运筹学与控制论

DOI:10.11721/cqnuj20190501

求解随机变分不等式问题的修正外梯度随机逼近算法*

张小娟，杜学武
(重庆师范大学 数学科学学院，重庆 401331)

摘要:【目的】研究求解随机变分不等式问题的基于外梯度的随机逼近算法。【方法】依据求解经典变分不等式问题的外梯度算法，给出求解随机变分不等式问题的修正外梯度随机逼近算法。【结果】在适当的假设下，证明了修正外梯度随机逼近算法具有全局收敛性，初步的数值试验结果表明算法具有有效性。【结论】修正外梯度随机逼近算法是对已有的外梯度随机逼近算法的进一步推广，并且可在更弱的假设下获得它们的全局收敛性结果。

关键词:随机变分不等式;随机逼近;外梯度算法;全局收敛性
中图分类号:O224; O221.5　　**文献标志码:**A　　**文章编号:**1672-6693(2019)05-0016-06

众所周知，变分不等式(VI)问题已被广泛应用于经济管理、供应链网络、交通运输、博弈论等领域[1-2]。在现实生活中，很多问题都会涉及随机因素。因此，近些年来许多研究者开始关注对随机变分不等式(SVI)问题的研究。本文考虑 SVI 问题:找 $x^*\in X$，使得:

$$(x-x^*)^{\mathrm{T}}F(x^*)\geqslant 0,\forall x\in X, \tag{1}$$

其中 $X\subset\mathbf{R}^n$ 是一个非空闭凸集，F 是 $\mathbf{R}^n\to\mathbf{R}^n$ 的一个映射，且 $F(x):=E[f(x,\xi)]$，这里 $\xi\in\Omega$ 是概率空间$(\Omega,\mathcal{F},P)$中的随机变量，$f(x,\xi)$是一个连续映射，E 表示数学期望。

求解问题(1)的常用方法主要有两类:样本均值逼近(SAA)[3]和随机逼近(SA)[4-9]。这两类方法均基于用样本点逼近原问题，但 SA 方法每次迭代所需的样本个数少，且数值表现优于 SAA 方法。因此，本文主要考虑求解问题(1)的 SA 方法。最近 Kannan 等人[7]提出了求解问题(1)的基于外梯度的 SA 方法，该方法在假设 F 伪单调加并且 F 有界或者 Lipschitz 连续的条件下依概率 1 收敛。Iusem 等人[8]采用每次迭代选取若干个样本点修正外梯度算法，获得了在 F 伪单调并且 Lipschitz 连续的条件下依概率 1 收敛的理论结果。进一步地，Iusem 等人[9]又在文献[8]的基础上对算法加入线搜索技术，获得了和文献[8]一样的理论收敛性和更好的收敛率结果。

本文受文献[7,9]中研究工作的启发，给出求解随机变分不等式问题(1)的修正外梯度随机逼近算法，在更弱的假设下，证明算法的全局收敛性，并对文献[10]中的算例进行数值试验，从而验证算法的有效性。

1 预备知识

定义 1[1]　设 X 是 $\mathbf{R}^n$ 中的非空闭凸集，$y\in\mathbf{R}^n$。定义 $\Pi_X[y]:=\arg\min\{\|x-y\|:x\in X\}$为点 y 在集合 X 上的投影。

定义 2[1]　对任意的 $x\in\mathbf{R}^n$ 和 $\alpha>0$，定义 $r_\alpha(x):=\|x-\Pi_X[x-\alpha F(x)]\|$ 为问题(1)的自然残差函数。当 $\alpha=1$ 时，自然残差函数简记为 $r(x)$。

引理 1[1]　x^* 是问题(1)的解，当且仅当对任意 $\alpha>0$ 有 $x^*=\Pi_X[x^*-\alpha F(x^*)]$。

引理 2[1]　给定 $x\in\mathbf{R}^n$，函数 $\alpha\longmapsto\dfrac{r_\alpha(x)}{\alpha}$在区间$(0,\infty)$上是不增的。

引理 3[1]　设 $X\subset\mathbf{R}^n$ 是一个非空闭凸集，则:1) 对任意的 $x,y\in\mathbf{R}^n$，有 $\|\Pi_X(x)-\Pi_X(y)\|\leqslant\|x-y\|$;2) 对任意的 $x\in\mathbf{R}^n$，$y\in X$，有$(x-\Pi_X(x))^{\mathrm{T}}(y-\Pi_X(x))\leqslant 0$;3) 对任意的 $x\in\mathbf{R}^n$，$y\in X$，$\|\Pi_X(x)-y\|^2\leqslant\|x-y\|^2-$

* **收稿日期:**2018-10-09　**修回日期:**2019-05-13　**网络出版时间:**2019-09-26 11:24
资助项目:重庆市自然科学基金(No. cstc2017jcyjA0788)
第一作者简介:张小娟，女，研究方向为数学规划理论与算法，E-mail:1143182877@qq.com;**通信作者:**杜学武，男，教授，博士，E-mail:508677034@qq.com
网络出版地址:http://kns.cnki.net/kcms/detail/50.1165.N.20190926.1124.036.html

$\| \Pi_X(x)-x \|^2$。

引理 4[11]　对任意的 $x,y\in \mathbf{R}^n$，$\lambda\in(0,1)$，有：

$$\| \lambda x+(1-\lambda)y \|^2=\lambda \| x \|^2+(1-\lambda) \| y \|^2-\lambda(1-\lambda) \| x-y \|^2$$

成立。

命题 1[12]　设$\{\nu_k\}$，$\{\omega_k\}$，$\{a_k\}$，$\{b_k\}$均为适应于σ代数流$\{\mathcal{F}_k\}$的可积非负随机序列，满足 $\sum\limits_{k=1}^{\infty}a_k<\infty$，$\sum\limits_{k=1}^{\infty}b_k<\infty$，并且对任意的$k\in\mathbf{N}$，$E[\nu_{k+1}|\mathcal{F}_k]\leqslant(1+a_k)\nu_k-\omega_k+b_k$ 依概率 1 成立，则$\{\nu_k\}$依概率 1 收敛，并且 $\sum\limits_{k=1}^{\infty}\omega_k<\infty$ 依概率 1 成立。

2 算法及其收敛性

求解经典 VI 问题的外梯度算法已经发展得比较成熟，但是，对于求解 SVI 问题的外梯度算法，人们研究得却很少。Censor 等人[13]给出了求解经典 VI 问题的一类修正外梯度算法，其中将 $k+1$ 步的迭代点取为第 k 步和矫正步的迭代点的凸组合。为了充分利用已有迭代点的信息以提高算法的有效性，本文结合文献[7，13]中的迭代格式和文献[9]中的线搜索策略，提出求解问题(1)的一类修正外梯度随机逼近(Modified extragradient stochastic approximation，MESA)算法。下面给出 MESA 算法的具体步骤。

MESA **算法**　初始步，选取初始点 $x^0\in\mathbf{R}^n$，$\gamma>0$，$\theta\in(0,1)$，$\mu\in\left(0,\frac{1}{\sqrt{3}}\right)$，令 $k:=0$。

步骤 1，给定 x^k，取来自 Ω 的样本点 ξ^k，若 $x^k=\Pi_X[x^k-f(x^k,\xi^k)]$则停止；否则转步聚 2。

步骤 2，对于 $\alpha\in(0,\infty)$，定义 $y^k(\alpha):=\Pi_X[x^k-\alpha f(x^k,\xi^k)]$。找 $\alpha_k=\gamma\theta^{l_k}$，其中 l_k 是使得下式立的最小非负整数：

$$\gamma\theta^{l_k} \| f(x^k,\xi^k)-f(y^k(\gamma\theta^{l_k}),\xi^k) \| \leqslant\mu \| x^k-y^k(\gamma\theta^{l_k}) \| 。\tag{2}$$

记 $y^k=y^k(\alpha_k)=\Pi_X[x^k-\alpha_k f(x^k,\xi^k)]$。产生 Ω 的样本点 η^k，计算 $t^k=\Pi_X[x^k-\alpha_k f(y^k,\eta^k)]$。选取 $\delta\in(0,1)$，计算：

$$x^{k+1}=\delta x^k+(1-\delta)t^k,\tag{3}$$

令 $k:=k+1$，转步聚 1。

MESA 算法中的 ξ^k 和 η^k 是来自随机变量 ξ 的独立同分布样本。下面考虑由 MESA 算法产生的序列$\{x^k\}$的收敛性。关于随机过程$\{x^k\}$的 σ 代数流为 $\mathcal{F}_k=\sigma(x^0,\xi^0,\cdots,\xi^{k-1},\eta^0,\cdots,\eta^{k-1})$，$\hat{\mathcal{F}}_k=\sigma(x^0,\xi^0,\cdots,\xi^{k-1},\xi^k\ \eta^0,\cdots,\eta^{k-1})$。定义 $\varepsilon_1^k:=f(x^k,\xi^k)-F(x^k)$，$\varepsilon_2^k:=f(y^k,\eta^k)-F(y^k)$，$\varepsilon_3^k:=f(y^k,\xi^k)-F(y^k)$是由 MESA 算法产生的随机误差。

为了获得 MESA 算法的收敛性，需要下面的假设。

假设 1　问题(1)的解集 X^* 非空。

假设 2　存在非负随机变量 $\mathcal{L}:\Omega\to\mathbf{R}_+$，满足 $E[\mathcal{L}(\xi)]<\infty$，$\mathcal{L}(\xi)\geqslant1$，$\forall\, x,y\in\mathbf{R}^n$，$\| f(x,\xi)-f(y,\xi) \| \leqslant \mathcal{L}(\xi) \| x-y \|$，依概率 1 成立。

假设 3　对任意的 $x^*\in X$，有$(x-x^*)^{\mathrm{T}}F(x)\geqslant0$，$\forall\, x\in X$。

假设 4　误差 $\varepsilon_1^k,\varepsilon_2^k,\varepsilon_3^k$ 满足：1) $E[\varepsilon_1^k|\mathcal{F}_k]=0$，$E[\varepsilon_2^k|\hat{\mathcal{F}}_k]=0$；2) $\sum\limits_{k=0}^{\infty}E[\| \varepsilon_1^k \|^2|\mathcal{F}_k]<\infty$，$\sum\limits_{k=0}^{\infty}E[\| \varepsilon_2^k \|^2|\mathcal{F}_k]<\infty$，$\sum\limits_{k=0}^{\infty}E[\| \varepsilon_3^k \|^2|\mathcal{F}_k]<\infty$成立。

假设 2 是文献[9]中的 Hölder 连续假设的一个特例，假设 3 弱于文献[7]中的伪单调加假设和文献[9]中的伪单调假设，假设 4 是 SA 方法中的基本假设，它要求随机误差具有无偏性和方差具有可控性。

为了叙述方便，定义 $A_k:=(1-3\mu^2)\gamma^2 \| \varepsilon_1^k \|^2+3\gamma^2 \| \varepsilon_2^k \|^2+3\gamma^2 \| \varepsilon_3^k \|^2$，$B_k:=\frac{1}{2}(1-\delta)(1-3\mu^2)\alpha_k^2$ 在给出 MESA 算法的收敛性之前，先证明引理 5 和引理 6。下面的引理 5 给出了与随机误差有关的一个递推关系。

引理 5　在假设 1～3 成立的条件下，若 MESA 算法在第 k 步不终止，则对任意 $x^*\in X^*$，由 MESA 算法产

生的序列$\{x^k\}$满足

$$\| x^{k+1}-x^* \|^2 \leqslant \| x^k-x^* \|^2 - B_k r^2(x^k)+A_k-2\alpha_k(1-\delta)(y^k-x^*)^{\mathrm{T}}\varepsilon_2^k。\tag{4}$$

证明 由引理 3 得到:

$$\| t^k-x^* \|^2 = \| \Pi_X[x^k-\alpha_k f(y^k,\eta^k)]-x^* \|^2 \leqslant$$
$$\| (x^k-\alpha_k f(y^k,\eta^k))-x^* \|^2 - \| \Pi_X[x^k-\alpha_k f(y^k,\eta^k)]-(x^k-\alpha_k f(y^k,\eta^k)) \|^2。$$

由假设 3 及 ε_2^k 的定义得到:

$$\| t^k-x^* \|^2 \leqslant \| x^k-x^* \|^2 - \| (x^k-y^k)-(t^k-y^k) \|^2 - 2\alpha_k(t^k-y^k)^{\mathrm{T}} f(y^k,\eta^k) - 2\alpha_k(y^k-x^*)^{\mathrm{T}}\varepsilon_2^k \leqslant$$
$$\| x^k-x^* \|^2 - \| x^k-y^k \|^2 - \| y^k-t^k \|^2 + 2\alpha_k(t^k-y^k)^{\mathrm{T}}(f(x^k,\xi^k)-f(y^k,\eta^k)) - 2\alpha_k(y^k-x^*)^{\mathrm{T}}\varepsilon_2^k。$$

通过均值不等式得:

$$\| t^k-x^* \|^2 \leqslant \| x^k-x^* \|^2 - \| x^k-y^k \|^2 - 2\alpha_k(y^k-x^*)^{\mathrm{T}}\varepsilon_2^k + \alpha_k^2 \| f(x^k,\xi^k)-f(y^k,\eta^k) \|^2。\tag{5}$$

由不等式 $\left(\sum_{i=1}^3 a_i\right)^2 \leqslant 3\sum_{i=1}^3 a_i^2$ 和(2)式以及 $\alpha_k\in(0,\gamma)$,得到:

$$\alpha_k^2 \| f(x^k,\xi^k)-f(y^k,\eta^k) \|^2 \leqslant$$
$$\alpha_k^2(\| f(x^k,\xi^k)-f(y^k,\xi^k) \| + \| f(y^k,\xi^k)-F(y^k) \| + \| F(y^k)-f(y^k,\eta^k) \|)^2 \leqslant$$
$$3\alpha_k^2 \| f(x^k,\xi^k)-f(y^k,\xi^k) \|^2 + 3\alpha_k^2 \| f(y^k,\eta^k)-F(y^k) \|^2 + 3\alpha_k^2 \| f(y^k,\xi^k)-F(y^k) \|^2。$$

因此:

$$\alpha_k^2 \| f(x^k,\xi^k)-f(y^k,\eta^k) \|^2 \leqslant 3\mu^2 \| x^k-y^k \|^2 + 3\gamma^2 \| \varepsilon_2^k \|^2 + 3\gamma^2 \| \varepsilon_3^k \|^2。\tag{6}$$

由 ε_1^k 得到 $y^k=\Pi_X[x^k-\alpha_k f(x^k,\xi^k)]=\Pi_X[x^k-\alpha_k(F(x^k)+\varepsilon_1^k)]$。根据引理 2、引理 3 以及 $\alpha_k\in(0,\gamma)$得:

$$\alpha_k^2 r^2(x^k) \leqslant r_{\alpha_k}^2(x^k) = \| x^k-\Pi_X[x^k-\alpha_k F(x^k)] \|^2 = \| (x^k-y^k)+(y^k-\Pi_X[x^k-\alpha_k F(x^k)]) \|^2 \leqslant$$
$$2\| x^k-y^k \|^2 + 2\| \Pi_X[x^k-\alpha_k(F(x^k)+\varepsilon_1^k)]-\Pi_X[x^k-\alpha_k F(x^k)] \|^2 \leqslant 2\| x^k-y^k \| + 2\gamma^2 \| \varepsilon_1^k \|^2。$$

因此:

$$\| x^k-y^k \|^2 \geqslant \frac{\alpha_k^2}{2} r^2(x^k) - \gamma^2 \| \varepsilon_1^k \|^2。\tag{7}$$

结合(6)~(7)式,得:

$$\| x^t-x^* \|^2 \leqslant \| x^k-x^* \|^2 - \frac{\alpha_k^2(1-3\mu^2)}{2} r(x^k)^2 - 2\alpha_k(y^k-x^*)^{\mathrm{T}}\varepsilon_2^k +$$
$$(1-3\mu^2)\gamma^2 \| \varepsilon_1^k \|^2 + 3\gamma^2 \| \varepsilon_2^k \|^2 + 3\gamma^2 \| \varepsilon_3^k \|^2。\tag{8}$$

由(3)式、(8)式、引理 4 和 $\delta\in(0,1)$,得到:

$$\| x^{k+1}-x^* \|^2 = \| \delta x^k+(1-\delta)t^k-x^* \|^2 = \| \delta(x^k-x^*)+(1-\delta)(t^k-x^*) \|^2 \leqslant$$
$$\delta\| x^k-x^* \|^2 + (1-\delta)\| t^k-x^* \|^2 \leqslant \| x^k-x^* \|^2 - \frac{\alpha_k^2(1-\delta)(1-3\mu^2)}{2} r(x^k)^2 - 2(1-\delta)\alpha_k(y^k-x^*)^{\mathrm{T}}\varepsilon_2^k +$$
$$(1-3\mu^2)\gamma^2 \| \varepsilon_1^k \|^2 + 3\gamma^2 \| \varepsilon_2^k \|^2 + 3\gamma^2 \| \varepsilon_3^k \|^2。$$

根据 A_k, B_k 定义有 $\| x^{k+1}-x^* \|^2 \leqslant \| x^k-x^* \|^2 - B_k r(x^k)^2 + A_k - 2(1-\delta)\alpha_k(y^k-x^*)^{\mathrm{T}}\varepsilon_2^k$ 成立。 证毕

因为线搜索步长 α_k 的取值与样本点有关,所以需要寻找 α_k 的下界。下面的引理 6 就表明,步长 α_k 的取值要么是 γ,要么是 Lipschitz 常数下界的一个无偏随机估计。引理 6 的证明参考了文献[9]中引理 4,引理 5 的证明过程。

引理 6 在假设 2 成立的条件下,若 MESA 算法在 k 步不终止,则 $\alpha_k \geqslant \min\left\{\frac{\mu\theta}{\mathcal{L}(\xi^k)},\gamma\right\}$依概率 1 成立,并且有 $E[\alpha_k|\mathcal{F}_k] \geqslant \frac{1}{L}\min\{\mu\theta,\gamma\}$成立,其中 $L := E[\mathcal{L}(\xi)]$。

证明 如果 $\alpha_k=\gamma$ 满足(2)式,则 $\alpha_k=\gamma$;如果 $\alpha_k=\gamma$ 不满足(2)式,则有 $\theta^{-1}\alpha_k \| f(x^k,\xi^k)-f(y^k(\theta^{-1}\alpha_k),\xi^k) \| > \mu \| x^k-y^k(\theta^{-1}\alpha_k) \|$。由假设 2 得:

$$\| f(x^k,\xi^k)-f(y^k(\theta^{-1}\alpha_k),\xi^k) \| \leqslant \mathcal{L}(\xi^k) \| x^k-y^k(\theta^{-1}\alpha_k) \|。$$

故可得到 $\alpha_k \geqslant \frac{\mu\theta}{\mathcal{L}(\xi^k)}$。通过上述讨论可知 $\alpha_k \geqslant \min\left\{\gamma,\frac{\mu\theta}{\mathcal{L}(\xi^k)}\right\}$依概率 1 成立。由假设 2 知,$\mathcal{L}(\xi^k)\geqslant 1$ 依概率 1 成立,从而有 $\alpha_k\mathcal{L}(\xi^k) \geqslant \min\{\mu\theta,\gamma\}$依概率 1 成立,故:

$$\min\{\mu\theta,\gamma\} \leqslant E[\alpha_k \mathcal{L}(\xi^k) \mid \mathcal{F}_k] \leqslant E[\alpha_k \mid \mathcal{F}_k] E[\mathcal{L}(\xi^k) \mid \mathcal{F}_k] = LE[\alpha_k \mid \mathcal{F}_k], E[\alpha_k \mid \mathcal{F}_k] \geqslant \frac{1}{L}\min\{\gamma,\theta\mu\}$$

成立。　　证毕

下面给出 MESA 算法的收敛性结果。

定理 1　在假设 1～4 成立的条件下，MESA 算法要么有限步终止（如第 k 步终止，则 x^k 为问题(1)的解），要么产生一个无穷序列$\{x^k\}$，满足：1) $\{x^k\}$依概率 1 有界；2) $\lim\limits_{k\to\infty} r^2(x^k)=0$；3) $\{x^k\}$的每个聚点$\bar{x}$均依概率 1 为问题(1)的解；4) $\lim\limits_{k\to\infty} d(x^k,X^*)=0$ 依概率 1 成立，其中 $d(\cdot,\cdot)$表示距离。

证明　若 MESA 算法在第 k 步终止，则有 $x^k=\Pi_X[x^k-\alpha_k f(x^k,\xi^k)]$，由引理 1 知：

$$(x-x^k)^{\mathrm{T}} f(x^k,\xi^k) \geqslant 0, \forall x\in X。\tag{9}$$

由 $x^k\in\mathcal{F}_k$，ξ^k 与 $\mathcal{F}_k$ 独立，得到 $E[f(x^k,\xi^k)\mid\mathcal{F}_k]=F(x^k)$。对(9)式取 $E[\cdot\mid\mathcal{F}_k]$，得到$(x-x^k)^{\mathrm{T}}F(x^k)\geqslant 0$，$\forall x\in X$，所以 $x^k\in X^*$。

若 MESA 算法不在有限步终止，则产生一个无穷序列$\{x^k\}$。由引理 6，$E[\cdot\mid\mathcal{F}_k]=E[E[\cdot\mid\hat{\mathcal{F}}_k]\mid\mathcal{F}_k]$和假设 4，并对(9)式取 $E[\cdot\mid\mathcal{F}_k]$，得到：

$$\begin{aligned}
E[\|x^{k+1}-x^*\|^2\mid\mathcal{F}_k] &\leqslant \|x^k-x^*\|^2+E[A_k\mid\mathcal{F}_k]-2(1-\delta)E[\alpha_k\mid\mathcal{F}_k]\cdot \\
&E[(y^k-x^*)\mid\mathcal{F}_k]^{\mathrm{T}}E[E[\varepsilon_2^k\mid\hat{\mathcal{F}}_k]\mid\mathcal{F}_k]-E[B_k\mid\mathcal{F}_k]r^2(x^k)= \\
&\|x^k-x^*\|^2-E[B_k\mid\mathcal{F}_k]r^2(x^k)+E[A_k\mid\mathcal{F}_k]+E[A_k\mid\mathcal{F}_k]= \\
&\|x^k-x^*\|^2-\frac{1}{2}(1-\delta)(1-3\mu^2)E[\alpha_k^2\mid\mathcal{F}_k]r^2(x^k)+E[A_k\mid\mathcal{F}_k]\leqslant \\
&\|x^k-x^*\|^2-\frac{1}{2L^2}(1-\delta)(1-3\mu^2)(\min\{\theta\mu,\gamma\})^2r^2(x^k)+E[A_k\mid\mathcal{F}_k]。
\end{aligned}$$

记 $\nu_k:=\|x^k-x^*\|^2$，$b_k:=E[A_k\mid\mathcal{F}_k]$，$\omega_k:=\frac{1}{2L^2}(1-\delta)(1-3\mu^2)$，并令 $a_k\equiv 0$，则显然有 $\sum\limits_{k=1}^{\infty}a_k<\infty$。因此有 $E[\nu_{k+1}\mid\mathcal{F}_k]\leqslant(1+a_k)\nu_k-\omega_k+b_k$。由 A_k 有：

$$b_k=E[A_k\mid\mathcal{F}_k]=(1-3\mu^2)\gamma^2E[\|\varepsilon_1^k\|^2\mid\mathcal{F}_k]+3\gamma^2E[\|\varepsilon_2^k\|^2\mid\mathcal{F}_k]+3\gamma^2E[\|\varepsilon_3^k\|^2\mid\mathcal{F}_k],$$

故由假设 4 可知 $\sum\limits_{k=1}^{\infty}b_k<\infty$。因此，由命题 1 知$\{\|x^k-x^*\|^2\}$依概率 1 收敛，且 $\sum\limits_{k=1}^{\infty}\omega_k<\infty$ 成立，故$\{x^k\}$依概率 1 有界，即结论 1)成立；

且 $\omega_k=\frac{1}{2L^2}(1-\delta)(1-3\mu^2)(\min\{\theta\mu,\gamma\})^2r^2(x^k)\to 0$ 成立，$\lim\limits_{k\to\infty}r^2(x^k)=\lim\limits_{k\to\infty}\|x^k-\Pi_X[x^k-F(x^k)]\|^2=0$ 成立，即结论 2)成立；

设$\bar{x}$为序列$\{x^k\}$的任一聚点，则由 F 和 Π 的连续性有$\bar{x}=\Pi_X[\bar{x}-F(\bar{x})]$，故由引理 1 知$\bar{x}\in X^*$，因此结论 3)成立；

由于$\{x^k\}$依概率 1 有界，并且它的任意聚点均依概率 1 为问题(1)的解，故$\lim\limits_{k\to\infty}d(x^k,X^*)=0$ 依概率 1 成立，结论 4)得证。　　证毕

3 数值试验

该小节选取文献[10]中的算例，对 MESA 算法给出初步的数值试验结果。算法的终止准则为 $\|x^k-\Pi_X[x^k-f(x^k,\xi^k)]\|\leqslant 10^{-6}$，各参数的取值分别为：$\theta=0.5$，$\gamma=0.6$，$\mu=0.4$，$\delta=0.8$。测试环境为 Matlab R2016b，操作系统为 Windows 10，CPU 为 1.8 GHz，内存为 8.0 GB。定义函数 $f_1(x,\xi)$，$f_2(x,\xi)$分别如下：

$$f_1(x,\xi):=\begin{pmatrix} x_1-\xi x_2+3-2\xi \\ -\xi x_1+2x_2+\xi x_3-2-\xi \\ \xi x_2+3x_3-3-\xi \end{pmatrix}, f_2(x,\xi):=\begin{pmatrix} x_1^2-\xi x_2+3-2\xi \\ -\xi x_1+2x_2^2+\xi x_3-2-\xi \\ \xi x_2+3x_3^2-3-\xi \end{pmatrix}。$$

算例 1ξ 是随机变量，$X=R_+^3$，样本函数取 $f_1(x,\xi)$；算例 2ξ 是随机变量，$X=R_+^3$，样本函数取 $f_2(x,\xi)$；算例 3ξ 是随机变量，$X=[0,4]\times[0,4]\times[0,4]$，样本函数取 $f_1(x,\xi)$；算例 4ξ 是随机变量，$X=[0,4]\times[0,4]\times[0,4]$，样本函数取 $f_2(x,\xi)$。

对于上述每个算例,任意给定 $\xi\in[0,1]$,它们都有唯一解 $x^*=(0,1,1)^{\mathrm{T}}$。表 1 给出了 MESA 算法、ESA 算法[9]和 SAA 算法[10]对算例 1～4 在随机变量 ξ 服从 $\Omega=[0,1]$上的均匀分布和 ξ 服从 $\Omega=(-\infty,+\infty)$上的均值为 1/2、方差为 1/12 的正态分布这两种情形下测试 100 次的平均值的数值试验结果。数值试验中,每次计算均选取初始点 $x^0=(0,0,0)^{\mathrm{T}}$。表 1 中的第 1 列至第 4 列分别给出了算例编号、算法名称、所得的近似解$\bar{x}$和计算(CPU)时间。

表 1 算例 1～4 的数值试验结果

Tab.1 Numerical results for example 1～4

算例	算法	随机变量 ξ 服从均匀分布		随机变量 ξ 服从正态分布	
		近似解$\bar{x}$	CPU 时间/s	近似解$\bar{x}$	CPU 时间/s
1	MESA	(0,0.999 999,1.000 002)	0.003 5	(0,0.999 992,1.000 004)	0.004 3
	ESA	(0,0.999 999,1.000 002)	0.005 5	(0,0.999 994,1.000 002)	0.006 5
	SAA	(0.000 061,0.999 746,0.999 523)	0.013 23	(0.000 543,0.994 255,0.958 932)	0.014 5
2	MESA	(0,0.999 996,1.000 005)	0.002 1	(0,0.999 993,1.000 004)	0.002 1
	ESA	(0,0.999 999,1.000 003)	0.004 3	(0,0.999 994,1.000 002)	0.004 3
	SAA	(0.000 002,1.006 939,1.009 293)	0.012 6	(0.000 102,1.000 232,1.009 234)	0.012 6
3	MESA	(0,0.999 992,1.000 003)	0.004 5	(0,0.999 9994,1.000 005)	0.004 5
	ESA	(0,0.999 993,1.000 002)	0.006 5	(0,0.999 9993,1.000 002)	0.006 5
	SAA	(0.006 147,0.999 463,0.999 231)	0.013 2	(0.000 034,0.996 461,0.999 431)	0.013 2
4	MESA	(0,0.999 993,1.000 002)	0.005 5	(0,0.999 999,1.000 002)	0.005 5
	ESA	(0,0.999 997,1.000 008)	0.007 5	(0,0.999 999 7,1.000 008)	0.007 5
	SAA	(0.006 147,0.946 354,0.923 166)	0.015 2	(0.000 042,0.999 143,0.999 341)	0.015 2

从表 1 可知,本文给出的 MESA 算法和文献[9]中的 ESA,文献[10]中的 SAA 算法对算例 1～4 均能在较短的时间内计算出十分接近问题真解 $x^*=(0,1,1)^{\mathrm{T}}$ 的近似解,并且 MESA 算法计算每个算例所花费的时间均少于 ESA 算法,且均明显少于 SAA 算法。因此,本文所提的 MESA 算法比 ESA 算法和 SAA 算法更为有效。

4 小结

本文给出了求解随机变分不等式问题的 MESA 算法。在较弱的假设下,证明了 MESA 算法产生的序列$\{x^k\}$的每一个聚点都是问题(1)的解依概率 1 成立。并且$\lim\limits_{k\to\infty}d(x^k,X^*)=0$ 和$\lim\limits_{k\to\infty}r^2(x^k)=0$ 都是依概率 1 成立。初步的数值试验结果表明 MESA 算法优于 ESA 算法和 SAA 算法。鉴于随机算法需要多次取样,接下来考虑修正算法框架,设计更为合适的终止准则,让算法自行多次取样,并给出相应的收敛性结果。

参考文献:

[1] FACCHINEI F,PANG J S. Finite dimensional variational inequalities and complementarity problems[M]. New York: Springer,2003.

[2] FERRIS M C,PANG J S. Engineering and economic applications of complementarity problems[J]. SIAM Review, 1997,39(4):669-713.

[3] XU H F. Sample average approximation methods for a class of stochastic variational inequality problems[J]. Asia-Pacific Journal of Operational Research,2010,27(01):103-119.

[4] JIANG J H,XU H F. Stochastic approximation approaches to the stochastic variational inequality problem[J]. IEEE Transactions on Automatic Control,2008,53(6):1462-1475.

[5] KOSHAL J,NEDIC A,SHANBHAG U V. Regularized iterative stochastic approximation methods for stochastic variational inequality problems[J]. IEEE Transactions on Automatic Control,2013,58(3):594-609.

[6] YOUSEFIAN F,NEDIC A,SHANBHAG U V. On smoothing,regularization,and averaging in stochastic approximation methods for stochastic variational inequality problems [J]. Mathematical Programming,2017,165(1):391-431.

[7] KANNAN A,SHANBHAG U V. The pesudomonotone stochastic variational inequality problem: analytical state-

ments and stochastic extragradient schemes[C]//American Control Conference (ACC), Portland, US, IEEE, 2014: 2930-2935.

[8] IUSEM A N, JOFRE A, OLIVEIRAl R I, et al. Extragradient method with variance reduction for stochastic variational inequalities [J]. SIAM Journal on Optimization, 2017, 27(2): 686-724.

[9] IUSEM A N, JOFRE A, OLIVEIRAl R I, et al. Variance-based extragradient methods with line search for stochastic variational inequalities[J]. SIAM Journal on Optimization, 2019, 29(1): 175-206.

[10] WANG M Z, LIN G H, GAO Y L, ALL M M. Sample average approximation method for a class of stochastic variational inequality problems [J]. Journal of Systems Science and Complexity, 2011, 24(6): 1143-1153.

[11] LOPEZ G, MARTIN V, XU H K. Perturbation techniques for nonexpansive mappings with applications [J]. Nonlinear Analysis Real World Applications, 2009, 10(4): 2369-2383.

[12] ROBBINS H, SIEGMUND D. A convergence theorem for nonnegative almost supermartingales and some applications[J]. Optimizing Methods in Statistics, 1971: 233-257.

[13] CENSOR Y, GIBALI A, REICH S. The subgradient extragradient method for solving variational inequalities in Hilbert space[J]. Journal of Optimization Theory and Applications, 2011, 148(2): 318-335.

Operations Research and Cybernetics

Modified Extragradient Stochastic Approximation Algorithms for Solving Stochastic Variational Inequality Problems

ZHANG Xiaojuan, DU Xuewu

(School of Mathematical Sciences, Chongqing Normal University, Chongqing 401331, China)

Abstract: [**Purposes**] The stochastic approximation algorithms based on extragradient for solving stochastic variational inequality (SVI) problems are studied. [**Methods**] A class of modified extragradient stochastic approximation (MESA) algorithms for solving SVI problemsis is presented in the light of the extragradient algorithm for solving the classical variational inequality problems. [**Findings**] Under appropriate assumptions, the global convergence of MESA algorithms are proved. The preliminary numerical results show that MESA algorithmsare is effective. [**Conclusions**] MESA algorithmsarethe generalizationsforsome existing extragradient stochastic approximation algorithms, and the proof of global convergence of MESA algorithms needs weaker assumptions.

Keywords: stochastic variational inequality; stochastic approximation; extragradient algorithm; global convergence

(责任编辑　陈　乔)